

EXHIBIT F

(Part 1)

ORIGINAL

Page 1

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF PENNSYLVANIA

CIVIL ACTION NO. 05-3524

HON. ROBERT F. KELLY

-----X

HEALTHCARE ADVOCATES, INC.,

Plaintiff,

vs.

HARDING, EARLEY, FOLLMER &
FRAILEY; JOHN F.A. EARLEY, III,
CHARLES L. RIDDLE, FRANK J.
BONINI, JR., KIMBERLY TITUS, and
JOHN DOES 1-5,

Defendants.

-----X

DEPOSITION OF EDWARD W. FELTEN

FRIDAY, FEBRUARY 2, 2007

DEPOLINK

Court Reporting & Litigation Support Services

Phone (973) 353-9880 Fax (973) 353-9445

www.depolinklegal.com

1 A P P E A R A N C E S:

2

3 McCARTER & ENGLISH, LLP

4 Four Gateway Center, 100 Mulberry Street

5 Newark, New Jersey 07102-4096

6 (973) 622-4444

7 BY: SCOTT S. CHRISTIE, ESQ.

8 Attorney for the Plaintiff

9

10 McKISSOCK & HOFFMAN, P.C.

11 105 East Evans Street

12 West Chester, Pennsylvania 19381

13 610-738-8850

14 BY: JEFFREY P. LEWIS, ESQ.

15 Attorney for the Defendants

16

17

18

19

20

21

22

23

24

25

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

EDWARD W. FELTEN

Computer Science Department, 35 Olden Street,
Princeton, New Jersey 08544, having been first
duly sworn, was examined and testified as follows:

DIRECT EXAMINATION

BY MR. CHRISTIE:

Q. Good morning, Mr. Felten. My name is
Scott Christie. I'm an attorney with McCarter &
English in Newark, New Jersey, and I represent
Healthcare Advocates, Inc., which is the plaintiff
in this lawsuit. You understand that you are here
to testify and be deposed in your capacity as an
expert witness in this case?

A. Yes.

Q. And you were retained as an expert.
Correct?

A. Yes.

Q. Do you understand that you're now under
oath?

A. Yes.

Q. And as a consequence of being under
oath, you're obliged to answer my questions
completely and truthfully?

A. Yes.

1 Q. Without belaboring in great detail all
2 of the text in your report, is what you have
3 written in here truthful and accurate? And the
4 reason I ask is because you haven't signed it
5 under penalty of perjury or in the form of an
6 affidavit. But if you're able to say that it is
7 accurate and truthful, that would help as opposed
8 to me going through specific topics ad nauseam.

9 A. Yes. That was my intention when I
10 prepared this, and it is accurate and truthful.
11 There is only one issue that I've learned about
12 since I -- since I submitted the report, and that
13 is at the very end.

14 Q. What's that?

15 A. It's in paragraph 69 where I say that "I
16 understand that as of September 1, 2006, it" --
17 meaning a particular convention "is also enshrined
18 in the Federal Rules of Evidence." And I
19 understand that that should say "instead of the
20 Federal Rules -- the Rules of Civil Procedure."

21 Q. Are you talking about the recent
22 amendments to the E-Discovery?

23 A. Yes, that's what I'm referring to.
24 That's the only -- that's the only issue in the
25 report that I've learned of since submitting it.

1 degree to which the Internet Archive Wayback
2 Machine respects the robots.txt exclusion
3 standard?

4 A. My understanding is that the Internet
5 Archive's crawlers look for the robots.txt file
6 and follow the -- any request in those files that
7 relate to the Internet Archive's crawlers. The
8 other issue is -- relates to how the Wayback
9 Machine responds to requests for pages -- for
10 historical pages if the -- if there is a
11 robots.txt file present on that site. And that's
12 something that's discussed at some length in my
13 report.

14 Q. Granted, that's true. And I believe you
15 say in your report -- and please correct me if I'm
16 wrong -- that by virtue of a properly drafted and
17 properly inserted robots.txt exclusion, Internet
18 Archive voluntarily agrees that it will not
19 publicly disclose any existing archive copies of
20 web content for the web site related to the robots
21 exclusion?

22 MR. LEWIS: Objection to form.

23 A. That's a pretty involved statement,
24 but --

25 Q. I'll gladly rephrase it.

1 A. Okay, please.

2 Q. It might be easier if I ask you and get
3 it in your words. What is your understanding of
4 the manner in which the Internet Archive Wayback
5 Machine treats existing archive web content when
6 it confronts a robots.txt exclusion?

7 A. You're asking about what they do now?

8 Q. First, I'd like to know your
9 understanding of what they did back in July of
10 2003 and whether that's changed, to your
11 knowledge?

12 A. As to what they did, that is, at least
13 partly, one of the issues under discussion here,
14 that is, there are questions about what happened,
15 how the Wayback Machine did work during the
16 relevant time in 2003 and -- but I don't know if
17 that's what you're asking about.

18 Q. Let me rephrase. Back in July of 2003,
19 what was the Internet Archive policy related to
20 its treatment of archive web content upon
21 confronting a robots.txt text string?

22 A. My understanding is that they said that
23 they did not serve out archived versions of a page
24 if there was a robots.txt file present for that
25 page, but that is not -- apparently not what they

1 actually did.

2 Q. Well, we'll get to that. We're talking
3 about their policy. So you're saying that the
4 policy was, in the course of an inquiry, if our
5 servers ping the web site and find a robots.txt
6 exclusion, that we at Internet Archive will not
7 make publicly available existing archive versions
8 of that web site. Is that what you're saying?
9 And if not, please ask -- please rephrase it.

10 THE WITNESS: Could I hear that again,
11 please?

12 (Counsel requests the reading of the
13 following testimony:

14 "QUESTION: So you're saying that the
15 policy was, in the course of an inquiry, if our
16 servers ping the web site and find a robots.txt
17 exclusion, that we at Internet Archive will not
18 make publicly available existing archive versions
19 of that web site. Is that what you're saying?
20 And if not, please rephrase it.")

21 MR. LEWIS: Objection to form.

22 A. What I'm saying is that my understanding
23 is at the time they said that they would not serve
24 out archive versions of a page if there was a
25 robots.txt file in place referring to that page.

1 Q. That begs further questions. Are we
2 talking about page-by-page decision made about
3 releasing archived web content, or are we talking
4 about a web-site-by-web-site determination? And
5 if you don't understand what I mean, I'll rephrase
6 it.

7 A. No. I understand the question. The --
8 my understanding is that, at the time, they would
9 apply that -- at the time they said they would
10 apply the requests in a robots.txt file in
11 relation to requests to archived versions of
12 pages, and so whether all pages on a site were
13 treated the same or not would have depended on
14 what was in the robots.txt file.

15 Q. Well, what if the robots.txt exclusion
16 covered the entire content of the web site, that
17 is, all pages? In your understanding, how would
18 Internet Archive have treated public dissemination
19 upon confronting a robots.txt exclusion under
20 those circumstances?

21 MR. LEWIS: Again, you're asking for
22 example what their policy was at the time?

23 MR. CHRISTIE: Yes.

24 A. According to what they said, they would
25 not have served out archived copies of pages from

1 that site.

2 Q. No pages. Correct?

3 A. Yes. That's according to what they
4 said.

5 Q. Let me show you what I'll mark as
6 Felten-6.

7 (Robots.txt file bearing Bates number
8 HCA 00001 is marked as Felten Exhibit 6 for
9 Identification.)

10 Q. Do you recognize that document,
11 Professor Felten?

12 A. It looks like a robots.txt file.

13 Q. Is it one that you've seen before in the
14 course of rendering expert services here?

15 A. Yes.

16 MR. CHRISTIE: I'll note for the record
17 that in the bottom right-hand corner there's a
18 designation, HCA 00001.

19 MR. LEWIS: And for the record, that's a
20 Bates stamp, and that's not part of the actual
21 document.

22 MR. CHRISTIE: Yes. I didn't mean to
23 imply it was.

24 Q. Professor Felten, do you understand this
25 document, F-6, to be the robots.txt string that

1 was present on the Healthcare Advocates web
2 servers during July of 2003?

3 MR. LEWIS: Objection to form.

4 A. I don't know.

5 Q. What do you understand this text string
6 to be?

7 A. This is -- this appears to be a version
8 of the robots.txt file from the Healthcare
9 Advocates site, but I can't say whether it is or
10 is not the same as what was present in July of
11 2003.

12 Q. You have no reason to dispute that it
13 is. Correct?

14 A. The evidence, I think, does not say
15 whether it is or is not.

16 Q. I'm asking you a more precise question.
17 Do you have any reason to dispute that this
18 robots.txt string was in place on the Healthcare
19 Advocates web server during July of 2003?

20 A. Well, beyond the lack of -- beyond there
21 not being evidence as to what exactly was in the
22 file on that date, no, but that's, I think, a
23 reason to object to drawing the conclusion that
24 this is what was in the file as of that date.

25 Q. I'm not asking for that. Please listen

1 to my question. I'm asking whether you have any
2 reason to dispute whether this text string was in
3 the robots.txt file of the Healthcare Advocates
4 web server in July of 2003?

5 A. Yes, I do.

6 Q. What is the basis for your disputing
7 that?

8 A. The -- the record does not show
9 specifically what was in that file as of that
10 date. And one of the explanations that might
11 explain what happened with respect to the disputed
12 accesses is that something else was in this file.

13 Q. But that's pure speculation on your
14 part, isn't it?

15 MR. LEWIS: Objection to form.

16 A. Nobody knows precisely what was in that
17 file on that date or precisely what happened with
18 those accesses.

19 Q. I'm not talking about and asking you
20 about what everyone else knew. I'm asking about
21 you rendering expert opinions in this case.

22 Do you have any reasons to disbelieve that
23 this was in place at the time?

24 A. Yes.

25 Q. What is the specific basis of knowledge,

1 apart from your speculation, that this robots.txt
2 file that you see in Felten-6 was not in place on
3 the Healthcare Advocates web server in the
4 robots.txt file during July of 2003?

5 MR. LEWIS: Objection to form.

6 A. Well, again, there is -- there's not
7 evidence that this, as opposed to something else,
8 was present in that file at that time. And given
9 the behavior of the Internet Archive site at the
10 relevant times, that -- that behavior calls into
11 doubt whether this was, in fact, the content of
12 the robots.txt file on that date.

13 Q. How does it call that into doubt?

14 A. Well, the Internet Archive says that
15 their systems were supposed to be programmed, or
16 they thought their systems were programmed to load
17 this, the robots.txt file, and to behave in
18 certain ways. And the evidence shows that
19 their -- the Wayback Machine did not behave in
20 that way, and so that's one indication that the
21 robots.txt file may very well have been different
22 at that time.

23 Q. But there could have been other reasons
24 that caused the Wayback Machine not to have
25 operated in accordance with Internet Archive

1 policy at that time. Is that correct?

2 MR. LEWIS: Objection to form.

3 A. There could have been, and I discussed
4 this some in my report.

5 Q. Just so that I understand it, you are
6 not saying definitively that any malfunction of
7 the Wayback Machine during July of 2003 was caused
8 by a failure to properly install a robots.txt
9 exclusion. Is that correct?

10 A. I cannot say that -- definitively that
11 that is what happened, that's right.

12 Q. With regard to your view of Mr. Lenky's
13 report, which I will show you at some later point,
14 do you take issue with his conclusion that the
15 "disputed accesses," as you term them, to the
16 archived Healthcare Advocates content on the
17 Internet Archive site were caused by
18 representatives of the Harding Earley law firm?

19 A. Well, if I'm to respond to what's in his
20 report, I'd like to see his report and see the
21 statement that he made.

22 Q. Let me show you what I'll mark as
23 Felten-7.

24 (Report by Gideon Lenky is marked as
25 Felten Exhibit 7 for Identification.)

1 Q. Specifically, I'm referring you to
2 page 3 and the top of page 4 of Mr. Lenky's
3 report, which, in essence, concludes that
4 computers located at the offices of Harding Earley
5 Follmer & Frailey, the defendant here, were
6 responsible for what you've termed the "disputed
7 accesses," I think was the phrase you used, to the
8 archived Healthcare Advocates web content. I'm
9 asking you whether you agree with or you dispute
10 that conclusion.

11 A. He says that certain accesses, which are
12 identified here by IP address and by a DNS
13 address, came from Harding Earley, and I have no
14 reason to dispute that. And these are accesses to
15 Internet Archive's site, just to be clear.

16 Q. Right. And specifically referring to
17 the internet protocol address and the DNS address,
18 which is contained within the second line of
19 page 3 of text. Correct?

20 A. Yes, yes, that's what I was referring
21 to.

22 Q. Let's take a look at his analysis and
23 conclusion with regard to the second question,
24 whether there was a text string properly installed
25 on the robots.txt file of the Healthcare Advocates

1 that could have existed and been that size, 36 to
2 38 bytes.

3 Q. Well, based upon your review of the
4 server logs, do you have any reason to dispute his
5 conclusion that Felten-6, number one, is in the
6 range of 36 to 38 bytes, again, based on your
7 review of the server logs?

8 A. Perhaps I don't understand the question
9 because I don't understand him as asserting that
10 this file that's in front of me here is
11 necessarily that size. As I said before, I don't
12 know whether Felten-6 is or is not the file that
13 was transmitted back then. And he doesn't refer
14 to Felten-6, so I don't know whether he's
15 referring to this document or not.

16 Q. He's not referring to the document, but
17 if you look, then, two-thirds of the way down on
18 page 4 of his report, do you see the robots text
19 string referenced there?

20 A. This is what he says was the content of
21 the robots.txt file as of the date of his report.

22 Q. Yes. And does it mirror what is
23 contained in Felten-6?

24 A. Yes.

25 Q. Based upon your analysis of the logs

1 that you conducted as an expert here, I believe
2 you mentioned in your report that you noted a
3 number of occasions where requests were made for
4 archived Healthcare Advocates web content during
5 the period July 9 through July 14 where those
6 requests were unsuccessful. Correct?

7 A. Yes.

8 Q. And I believe you also came to the
9 conclusion that many, if not all, of those
10 unsuccessful requests were due to a robots.txt
11 exclusion being in place.

12 A. What are you referring to here?

13 Q. I'm talking about attempts to access
14 archived Healthcare Advocates material from
15 Internet Archive during July of '03 having been
16 unsuccessful.

17 A. Yes, some requests at that time were
18 unsuccessful.

19 Q. And specifically, you referred to the
20 fact that there were particular servers of
21 Internet Archive that appear not to have been
22 working accurately with regard to fetching and
23 serving up the robots.txt exclusion. Correct?

24 MR. LEWIS: What page are you on?

25 MR. CHRISTIE: I'm looking for it now.

1 MR. LEWIS: Okay. Fair enough.

2 Q. For example, paragraph 45, you say,
3 "Based on descriptions available on the Internet
4 Archive's web site, one might have expected the
5 Wayback Machine to refrain from delivering the
6 requested pages. Nevertheless, the Wayback
7 Machine did deliver the pages."

8 You're referring to the robots.txt exclusion.
9 Correct?

10 A. Well, partly. Whether the pages would
11 be delivered would depend, in part, on what was in
12 the robots.txt file.

13 Q. I understand that. I'm asking, I think,
14 a separate question, whether there was a
15 robots.txt file of some sort present on the web
16 site during the relevant period, being July of
17 '03. Are you disputing that there was some sort
18 of robots file?

19 A. There was a robots.txt file present for
20 at least part of that period.

21 Q. What part of that period?

22 A. Well, we have server logs that show some
23 successful accesses to the robots.txt file. And
24 when we see those entries in the server logs, that
25 tells us that there was a robots.txt file of a

1 certain size present at that time. And it, of
2 course, doesn't tell us what was in the file other
3 than what size it was, beyond what size it was.

4 Q. But if you look at the logs and you
5 analyze them and the effect of the requests, do
6 you not see that there are denials based on robots
7 exclusions, dozens and hundreds of them throughout
8 the logs?

9 MR. LEWIS: Objection to form.

10 A. During the relevant period there were
11 some requests that were denied, yes.

12 Q. And they were denied based upon the
13 robots exclusion, as you can tell specifically
14 from the text of the logs. Right?

15 A. You can see from the logs that the page
16 that -- that's called the "query exclusion page"
17 in my report was delivered in some cases, but not
18 all.

19 Q. Okay. But that's not the question I
20 asked you. The question I asked you is: Based
21 upon your review of the logs, isn't it clear that
22 on dozens and perhaps hundreds of occasions access
23 to the archived content was denied based upon the
24 existence of a robots exclusion?

25 MR. LEWIS: Objection to the form.

1 A. The logs say that this query exclusion
2 page was delivered on a number of occasions during
3 that period.

4 Q. Based upon your knowledge of the
5 Internet Archive and how it works, that page would
6 have been delivered if a robots text file had been
7 in effect at the time. Is that correct?

8 MR. LEWIS: Objection to form.

9 A. Had there been a robots.txt file of a
10 certain form at that time, according to what
11 Internet Archive said that this is the page about
12 how their system worked, this is the page that
13 would have been delivered.

14 Q. When you say, "This is the page," you're
15 referring to --

16 A. I mean the query exclusion page. And
17 according to what they said, if there was not a
18 robots.txt file, or if it was somehow malformed or
19 not delivered correctly, then the query exclusion
20 page would not be delivered.

21 Q. And when you say a "query exclusion
22 page," you're referring to Bonini-8?

23 A. Yes, I'm referring to a page like
24 Bonini-8.

25 Q. But isn't it also true that if the

1 robots text string was properly formed and
2 inserted, any requester for the associated web
3 pages would have received a denial screen like
4 Bonini-8?

5 MR. LEWIS: Objection to form.

6 A. According to what Internet Archive said
7 at the time, this is -- this is what they said
8 would happen if there were a robots.txt file with
9 that particular content in it at the time. That's
10 what they said.

11 Q. Okay. And this exclusion page -- well,
12 actually, maybe we should save that for after
13 lunch. I want to ask a few additional questions
14 before we break for lunch.

15 Is it accurate to say, Professor Felten, that
16 you have no knowledge about whether the robots
17 exclusion on the Healthcare Advocates web site
18 during the period of July of '03 was malformed in
19 any way?

20 A. I don't know whether it was or not. The
21 available evidence does not allow us to answer
22 that question.

23 Q. Did you have reason to consult in the
24 course of your duties as an expert here the
25 deposition testimony of any of the representatives

1 robots.txt file, they will voluntarily change how
2 they treat archived copies of pages.

3 Q. Meaning they will deny access to
4 third-party requesters?

5 A. They have said that, yes.

6 Q. In fact, isn't it true from reviewing
7 the logs, Bonini-6 and Mohr-2, that you've noticed
8 a number of instances where that has, in fact,
9 taken place, meaning that a request has been made,
10 a robots exclusion has been identified, and no
11 content has been produced?

12 A. There are places in the logs where a
13 request is made and the Wayback Machine provides
14 the so-called "query exclusion pages" rather than
15 providing the requested content.

16 Q. Which is Bonini-8?

17 A. Yes, the query exclusion pages is what
18 looks like Bonini-8.

19 Q. Do you agree with the proposition, based
20 upon your analysis, that such a page, Bonini-8,
21 would only be served up by Internet Archive if
22 there is a properly configured and installed
23 robots.txt exclusion on the web site?

24 MR. LEWIS: Objection to form.

25 A. Not necessarily.

1 Q. Would have under all circumstances.

2 A. I don't know whether it would have under
3 all circumstances.

4 Q. What circumstances are you aware of that
5 it may not have been served up?

6 A. Well, if we -- if we look at the
7 available evidence about the accesses that are at
8 issue here, we don't know whether -- we don't know
9 whether the -- whether, at the time those accesses
10 were made, the Wayback Machine saw a robots.txt
11 file or -- and if it did see a robots.txt file,
12 exactly what was in it. And we see some requests
13 being satisfied. And so based on that, we can't
14 say whether the situation you described would
15 always happen -- would always have happened back
16 then.

17 Q. If a properly configured robots.txt
18 file -- and when I say "properly configured," I
19 mean properly drafted to exclude the Internet
20 Archive crawler -- had been captured by Internet
21 Archive and cached, would you agree that in
22 response to an inquiry for archive content related
23 to that robots file, Internet Archive served up
24 the denial screen, Bonini-8, during July of '03?

25 MR. LEWIS: Objection to form.

1 A. Well, we know that Internet Archive did
2 serve up the query exclusion page in some cases
3 during July of 2003. And it may be that in some
4 of those cases there was a cached robots.txt file
5 of the form you described. So it seems likely
6 that that would have happened at least sometimes.

7 Q. Are you aware of any instances in July
8 of '03, based upon your analysis, where the access
9 denial screen Bonini-8, was served up when there
10 wasn't a properly configured robots file in the
11 Internet Archive cache?

12 A. I don't know whether that happened or
13 not.

14 Q. Let me turn your attention to paragraph
15 35 of your report. You mention that "The requests
16 in a robots.txt file applied, by definition, only
17 to crawlers. They do not apply to people."

18 What do you mean by that?

19 A. Well, there's a distinction between a
20 crawler, which is a computer program -- and I
21 describe crawlers elsewhere in my report -- versus
22 people. And the document which I call here "The
23 Standard for Robot Exclusion" in my report talks
24 about how robots.txt files can be written to apply
25 to different crawlers, but there -- but the

1 robots.txt rules apply only to crawlers. There's
2 nothing in that system to make statements about
3 access by human beings.

4 Q. You understand this case involves
5 strictly the Internet Archive crawler. Correct?

6 MR. LEWIS: Objection to form.

7 A. That's the only crawler that is
8 involved, but there are, of course, people
9 involved as well.

10 Q. Do you understand that any part of this
11 lawsuit has anything to do with a robots.text
12 exclusion drafted specifically to exclude any
13 representatives of the Harding Earley law firm?

14 A. No, no robots.txt file could be drafted
15 to exclude members of the law firm.

16 MR. LEWIS: You mean that specifically?

17 A. I mean that specifically in the sense
18 that robots.txt files do not talk about people.
19 They apply to crawlers.

20 Q. They interact with crawlers and, in some
21 cases, direct crawlers' activity. Correct?

22 A. Crawlers sometimes read them and
23 sometimes alter their behavior based on what's in
24 the files.

25 Q. You mentioned that, as part of your

1 Q. As far as you know, that's accurate?

2 A. As far as I know.

3 Q. Let me turn your attention to
4 Mr. Lenky's report, specifically -- I'm sorry.
5 What's the number on that?

6 A. The Felten-7.

7 Q. -- Felten-7, specifically, paragraphs 5
8 and 6. Paragraph 5 --

9 A. Could I have a minute to review this?

10 Q. Oh, yeah, of course.

11 A. Okay.

12 Q. First, focusing your attention on
13 paragraph 5, do you see in the first sentence in
14 the answer to that question, it mentions that
15 "Based upon a reasonable degree of technical
16 certainty, the number of requests for protected
17 HCA content originating from defendant's IP was
18 549 on July 9, '03, and 118 on July 14, of '03."

19 Do you see that portion of the report?

20 A. Yes, I see that.

21 Q. Do you have any reason to dispute those
22 numbers?

23 A. These numbers are at least in the right
24 ballpark.

25 Q. You have no specific evidence to dispute

1 their accuracy. Correct?

2 A. I made some effort to replicate these
3 numbers and came out with results that were close
4 to this, so I could say they're in the right
5 ballpark.

6 Q. Let's look at paragraph 6, and I'll ask
7 you the same question about the numbers that
8 Mr. Lenky concludes in that paragraph. Please
9 take a minute and review it if you'd like.

10 A. Okay.

11 Q. The last sentence of his answer in
12 paragraph 6, do you dispute that conclusion?

13 A. These numbers, again, are at least in
14 the right ballpark at least. And I should point
15 out just to clarify that the questions that
16 Mr. Lenky is answering here refer to unauthorized
17 access. And without -- and if so -- in my answer
18 I'm setting aside the question of whether the
19 accesses were unauthorized or authorized as
20 opposed to successfully or unsuccessfully --

21 Q. I understand. Right. So we're talking
22 about whether they were, authorized or not,
23 attempted accesses to the archived content and
24 successful accesses in both paragraphs 5 and 6.
25 Right?

1 A. Yes.

2 Q. Turning your attention to paragraph 4 of
3 his report, let me refer you to the bottom of page
4 9 where he says, right before the end of that
5 page, "Based on a reasonable degree of technical
6 certainty, the number of times the defendant(s)"
7 -- 's' being in parentheses -- "viewed the
8 robots.txt exclusion pages was 498 on 9 July 2003,
9 35 on 10 July 2003, 7 on 11 July 2003, and 104 on
10 14 July 2003."

11 Do you dispute that conclusion,
12 Professor Felten?

13 A. Let me take a minute to review this.

14 Q. Sure.

15 A. I'm not certain whether these numbers
16 are correct.

17 Q. Okay. Do you have any reason to dispute
18 their accuracy?

19 A. Yes.

20 Q. What is the basis for your disputing
21 their accuracy?

22 A. Mr. Lenky's making an assumption here
23 that any file that is between 4,500 and 5,000
24 bytes in size and doesn't meet certain other
25 criteria must be an instance of the exclusion

1 page, and I'm not certain that's correct. There
2 could well have been other pages present which
3 were of that size, and, if so, then Mr. Lenky's
4 results is wrong. And he doesn't appear to
5 account for that possibility.

6 Q. But assuming that that is accurate, that
7 pages between 4500 and 5,000-dollar -- sorry --
8 4500 and 5,000 bytes in size and considering the
9 other criteria that he mentioned, all that is
10 accurate, would you then agree with his
11 conclusion?

12 A. You've got to be more precise about what
13 you're asking me to assume --

14 Q. Okay.

15 A. -- because there was an "all that."

16 Q. Well, you weren't entirely clear
17 yourself. I believe you mentioned that you had a
18 problem with the fact that he based his
19 conclusion, in part, on the fact that the robots
20 exclusion page, Bonini-8, is between 4500 and
21 5,000 bytes in size, plus there were other
22 criteria that he used to exclude certain other
23 pages, which may be in that same range. So I'm
24 asking you: Assuming that the facts he relied
25 upon to reach that conclusion in that paragraph,

1 which ends with the conclusion that we're talking
2 about, assuming those facts are accurate, would
3 you then agree with his conclusion?

4 A. No, not necessarily because he does not
5 state that there are no other files of that size
6 in the log. That appears to be an assumption in
7 his calculation which he's not stating here. So
8 based just on what he says here, you can't reach,
9 I think, the conclusion that he reaches.

10 Q. Assuming he got that information from an
11 authoritative source like Gordon Mohr, the chief
12 technologist at Internet Archive, would that lead
13 you to accept his conclusion?

14 A. Well, if the information is whether
15 there could have been other pages on the Internet
16 Archive site between 4,500 and 5000 bytes in size,
17 I'd be very skeptical that Mr. Mohr could remember
18 that now, whether there were any pages present
19 within that size range, which is a relatively
20 common size range for web sites. And Mr. Lenky
21 does not state that there were not other pages of
22 that size, let alone giving any basis for
23 concluding that.

24 Q. You don't know, for example, whether
25 Mr. Mohr, contemporaneous with the July 2003